

SPSS helps - Lab 8

For this lab you will be entering two dependent measures in order to find the degree to which they are correlated. Then you will use the information that was generated to make a prediction.

Entering data:

For this lab you will need to define two variables, one for the dependent measure of average cost of health care/person by state and a second for the degree to which each states average cost/person in nursing home care. The two dependent variables can be defined as numeric with no decimal places.

Once the variables are defined you should input the data. All of the data for the total health care costs measure should be entered in one column. The corresponding nursing home cost should be entered in the next column. Make sure to keep the data values paired. In other words, make sure that the total cost value for a particular state is in the same row as the nursing home cost value for that same state. This is very important since correlations are based on the relationship that exists between the X and Y values.

Analyzing the Data:

Begin the analysis by going to the **analyze** pull down menu. Select **regression** from the pull down list and **linear** from the side menu. A dialogue box will appear that will allow you to identify the dependent and independent variables. Place the variable for *total health care costs* in the **dependent variable** slot. Place the variable for *nursing home costs* in the **independent variable** slot.

Creating a scatter plot:

In order to see the data in a scatter plot you will need to click on the **plots** button. A new dialogue box will appear that will again have a list of variables and slots for identifying the X and Y variables. Place the word "dependent" as the Y variable and "ZPRED" (this represents the Z score equivalent of the nursing home costs) as the X

variable. Click on the **continue** and then on the **ok** button.

Reading the output:

The output will have, as always, more information than you need. First, you will find a chart that begins with "r". This value is the value of the correlation coefficient. Next to this is r squared which is the measure of the strength of the correlation. If you scroll down to the place that says "coefficients" you will find information for defining the line of best fit. The values for the slope and intercept (b and a) are found under the column marked "B". The value for the slope is in that column next to the variable name. The intercept is in the column marked "B" next to the label "constant".

Lab due: Wed. Dec. 9th. Please include a printout of the data and the scatter plot.